Contents

• Motivation
• ARCH
• GARCH
• Extensions: TARCH
Motivation

• Natural relation between return and volatility (risk), CAPM

• Pricing of options based on the volatility of asset returns
How to measure volatility?

• Historical volatility
  » Sample standard deviation of returns over a short period
  » What is the right period? This is chosen rather arbitrarily
  » Equal weights to yesterday and before 3-weeks observations?
    » This is historical volatility, but we need expected volatility
    » Volatility is not time-varying, this is at odds with stylized facts

• Exponentially weighted moving average models
  » Greater weight to yesterday than before 3-weeks observations, weight exponentially declining
    » But still rather arbitrary and backward-looking

• GARCH
Stylized facts in finance

• Unpredictability
  • Efficient markets

• Volatility
  • Time-varying (turbulent vs. tranquil periods)
  • Volatility clustering
    » When volatility is high, likely to remain high
    » Clustering of information arrivals, price discovery

• Fat tails
  • Extremely large returns (both positive and negative) more likely than standard normal distribution
ARCH – GARCH family

• ARCH = autoregressive conditional heteroscedasticity
• GARCH = generalized autoregressive conditional heteroscedasticity
• Dozens of GARCH models, for example:
  • EGARCH – exponential GARCH,
  • FIGARCH – fractionally integrated GARCH,
  • TARCH – threshold GARCH
Introduction

• Analysis and forecasting of the size of errors and its development over time
• Squared residuals often autocorrelated even though residuals themselves are not
• ARCH/GARCH for modelling volatility of asset – wide application in finance
• OLS assumes homoscedasticity, here we analyze heteroscedasticity, namely conditional heteroscedasticity (conditional on past information, as opposed to unconditional heteroscedasticity)
Introduction

• What does GARCH do?

• Imagine you own some asset and would like to know the tomorrow volatility of asset to evaluate its risk.

• You may arbitrarily take last X observations (say 22 days) and take standard deviation. In this case, you assign the same weight to yesterday’s observation and observation before three weeks, which is unrealistic.

• Alternatively, use GARCH.
• GARCH ‘says’ that best predictor of variance for today is a weighted average of:
  • the long-run variance (constant),
  • the forecast made in previous period and,
  • the new information that was not available when previous forecast was made (captured by the most recent squared residual)
So far concerned about long run statistics - notes

• Long-run moments of series – remember for example as we showed that forecast for ARMA model converges to its long-run mean, but you might be interested, what is mean likely to be in the next period

• Useful to distinguish between unconditional and conditional moments of series, unconditional - those of long-run, conditional – those of actual values of series in previous periods
• Assume that a stationary series of interest, $y_t$, follows simply:

$$y_t = \rho y_{t-1} + u_t$$

• Unconditional mean is $E(y_t) = 0$

• Unconditional variance is $\text{Var}(y_t) = \sigma^2 / (1 - \rho^2)$

• Conditional mean is $E(y_t | \Omega_{t-1}) = \rho y_{t-1}$

• Conditional variance $\text{Var}(y_t | \Omega_{t-1}) = \sigma_t^2$

• Where $\Omega_{t-1}$ is information set, e.g. all what is known before $t$

• Note that conditional mean is time-varying and conditional variance is *smaller* than unconditional variance

• $y_t = \rho y_{t-1} + u_t$ with such equation you already know how to predict mean, how about predicting variance of $y_t$?
ARCH

- Note that \( E(\gamma_t)^2 = E(\rho \gamma_{t-1} + u_t - E(\gamma_t))^2 = E(\rho \gamma_{t-1} + u_t - \rho \gamma_{t-1})^2 = E(u_t)^2 \)

  \( \Rightarrow \) for predicting \( \text{var}(\gamma_t) \) it is enough to predict \( \text{var}(u_t) \)!

- Engle (1982) proposed following model for the error term \( u_t \): \( u_t = e_t(\alpha_0 + \alpha_1 \cdot u_{t-1}^2)^{1/2} \)

- Assume also that \( \alpha_0 > 0 \) and \( 0 < \alpha_1 < 1 \), \( e_t \sim \text{IID}(0,1) \)
ARCH

• Note that conditional $E(u_t \mid \Omega_{t-1}) = 0$ and conditional $\text{var}(u_t \mid \Omega_{t-1}) = E(u_t^2 \mid \Omega_{t-1}) = a_0 + a_1 u_{t-1}^2$ — so conditional variance is not constant over time!

• Now, going back to our initial model $y_t = \varphi y_{t-1} + u_t$

• So, conditional mean of $y_t$ is $\varphi y_{t-1}$ and conditional variance of $y_t$ is $E(y_t)^2 = E(u_t)^2 = a_0 + a_1 u_{t-1}^2$
ARCH

• We just derived ARCH(1) model
• ARCH says that variability of $y_t$ in the next period is a function of this period squared residual and constant
• After estimating $\alpha_0$ and $\alpha_1$, we can predict the variance for the next period
  • ARCH straightforward to apply in empirical finance
    – Basic pricing models in finance: if you know that $y_t$ and $x_t$ provides same return, but you predict greater volatility in the next period for $x_t$, which one are you going to hold?
ARCH – “2 equations model”

• Mean equation (example – AR(1)):
  \[ y_t = \rho y_{t-1} + u_t \]

• Variance equation:
  \[ \mathbb{E}(u_t)^2 = a_0 + a_1 u_{t-1}^2 \]

• This is to be estimated jointly by maximum likelihood

• In principle, one can also estimate it equation by equation by OLS, but this is less efficient

• Again, all specification checks to be addressed like in ARIMA modeling
  • Such as are residuals white noise?
ARCH(q)

• ARCH (1) simply generalizes into ARCH (q), q specifies the number of lagged squared error terms

• \( u_t = e_t(\alpha_0 + \alpha_1 u_{t-1}^2 + \ldots + \alpha_q u_{t-q}^2)^{1/2} \)

• Then, \( \text{Var}(y_t) = \alpha_0 + \alpha_1 u_{t-1}^2 + \ldots + \alpha_q u_{t-q}^2 \)

• You may include dummies or any variable into this equation:
  • say day of the week dummy, deviation of exchange rate from central parity, etc.
Problem with ARCH

- Sometimes large number of squared lagged residuals must be included to specify the model correctly.

- Bolerslev (1986) extends ARCH model to allow more flexible lag structure – he introduces GARCH, it is to some extent analogy to ARMA models – ARCH is like MA model and now we introduce ‘AR’
GARCH (1,1)

• Note that $E(y_t)^2 = E(u_t)^2 = a_0 + a_1 u_{t-1}^2$

• Include lags of $E(u_t)^2$ to the RHS of the equation $E(u_t)^2 = a_0 + a_1 u_{t-1}^2 + \beta_1 E(u_{t-1})^2$

• Typically, we write it this way:

• $h_t = a_0 + a_1 u_{t-1}^2 + \beta_1 h_{t-1}$

• So now, what is the ‘best’ prediction for the next period variance of $y_t$? Weighted average (with the weights in brackets) of:
  
  – long-term variance ($a_0$),
  
  – this period actual variance – you may call it new information, not captured anywhere else ($a_1$),
  
  – the variance predicted for this period ($\beta_1$)
GARCH \((p,q)\)

• Naturally, you may generalize GARCH(1,1) to GARCH\((p,q)\) assuming this model for residual

\[ u_t = e_t(a_0 + a_1 * u_{t-1}^2 + \ldots + a_q * u_{t-q}^2 + \beta_1 * h_{t-1} + \ldots + \beta_p * h_{t-p})^{1/2} \]

• Note that \(e_t \sim \text{IID}(0,1)\), \(a_0 > 0\), \(a_1 \geq 0, \ldots, a_q \geq 0, q > 0, p \geq 0, \beta_1 \geq 0, \ldots, \beta_p \geq 0\)

• GARCH \((p,q)\) is stationary, if the sum of \(a\)'s all and \(\beta\)'s is strictly smaller than 1

• Conditional variance \(E(y_t)^2 = E(u_t)^2 = h_t\)

\[ = a_0 + a_1 * u_{t-1}^2 + \ldots + a_q * u_{t-q}^2 + \beta_1 * h_{t-1} + \ldots + \beta_p * h_{t-p} \]
How to estimate GARCH - strategy

1. Fit correct ARIMA model for $y_t$, and check if residuals are white noise etc.

2. If you suspect (G)ARCH residuals, estimate $u_t^2 = a_0 + a_1 u_{t-1}^2 + \ldots + a_q u_{t-q}^2 + v_t$ (simply regress residual on its lags, test significance, ARCH-LM test)

3. If ARCH errors not present, all explanatory variables should be jointly insignificant from zero

4. If jointly significant, this tells you that is some ARCH or GARCH structure of residuals

5. Examine PACF and ACF of squared residuals, reasoning same as for ARIMA modelling, (but note that very often you find that GARCH(1,1) fits the data best)

6. Evaluate the model:
   - Are residuals white noise?
Extensions: TARCH

• Leverage effect in finance:
  • Bad news are more important than good news for the behavior of stock
  • In our case, bad news = negative residual \((u_t<0)\)
  • Good news = positive residual \((u_t>0)\)

• TARCH accounts for this effect
  • It includes “dummy*bad news”

• Thus, TARCH (1,1):
  \[
  y_t = \rho y_{t-1} + u_t
  \]
  \[
  h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \gamma_1 I_{t-1} u_{t-1}^2
  \]
  • Where \(I_t=1\), if \(u_t<0\) and 0 otherwise

• GARCH is thus special case of TARCH
• If \(\gamma_1=0\), no asymmetric effects (“GARCH=TARCH”)
Readings

• Theory:

• Applications:

SEE THE COURSE WEBSITE FOR SOME OF THESE READINGS
• GARCH models are a bit more advanced and are not covered in every introductory econometrics textbook. All financial econometrics textbooks cover it, Brooks (2008) *Introductory Econometrics for Finance* is recommended
The Volatility Laboratory of R. Engle available at http://vlab.stern.nyu.edu/